

— ESTADÍSTICA UNIDIMENSIONAL —

- **Población:** conjunto de todos los individuos implicados en un estudio estadístico.
- **Muestra:** subconjunto de la población sobre el que se realiza el estudio. El número de individuos que componen la muestra se denomina **tamaño de la muestra** (N).
- **Variable:** es la característica que se estudia. Puede ser **cuantitativa** (si se puede medir con números) o **cualitativa** (si no se puede expresar mediante números).

Tablas de frecuencia

Son tablas que nos permiten ordenar los datos recogidos en nuestro estudio estadístico.

En una tabla de frecuencias aparecen distintos valores que debemos conocer:

x_i : posibles valores que toma la variable, X .

f_i (frecuencia absoluta): es el número de veces que la variable toma el valor x_i .

F_i (frecuencia absoluta acumulada): suma de las f_i de los valores menores o iguales que x_i .

h_i (frecuencia relativa): es el cociente entre f_i y el tamaño de la muestra (N).

H_i (frecuencia relativa acumulada): suma de las h_i de los valores menores o iguales que x_i .

x_i	f_i	F_i	h_i	H_i	$x_i \cdot f_i$	$x_i^2 \cdot f_i$
		Último valor de $F_i = N$		Último valor de $H_i = 1$		
	Suma de los valores = N		Suma de los valores = 1		$\sum x_i \cdot f_i$ (Suma de la columna)	$\sum x_i^2 \cdot f_i$ (Suma de la columna)

Tabla de frecuencias

Columnas de ayuda para el cálculo de la media, la varianza y la desviación típica.

Parámetros estadísticos más usuales

Media

Es una medida de centralización. Cuando existen valores extremos no es representativa del conjunto de datos.

$$\bar{x} = \frac{\sum x_i f_i}{N}$$

Varianza

Es una medida de dispersión. Nunca es negativa.

$$\sigma^2 = \frac{\sum x_i^2 \cdot f_i}{N} - \bar{x}^2$$

Desviación típica

Es la raíz cuadrada positiva de la varianza.

$$\sigma = \sqrt{\sigma^2}$$

Moda

Es el valor o valores de la variable que más se repiten, es decir, aquellos que tienen mayor f_i . Se representa por Mo.



Mediana

Valor de la variable que ocupa el centro de la distribución de datos cuando están ordenados de menor a mayor.

Se obtiene “buscando” en la columna de H_i el primer valor de la variable con frecuencia relativa acumulada mayor que 0'5. Si algún valor de x_i tiene $H_i = 0'5$, la mediana será la media de dicho valor y el siguiente.

Se representa por Me.

Cuartiles

Son aquellos valores de la variable que dividen el conjunto de datos en cuatro partes iguales.

El primer cuartil, Q_1 se obtiene “buscando” en la columna de H_i el primer valor de la variable con frecuencia relativa acumulada mayor que 0'25. Si algún valor de x_i tiene $H_i = 0'25$, Q_1 será la media de dicho valor y el siguiente.

El segundo cuartil, Q_2 , coincide con la mediana.

El tercer cuartil, Q_3 se obtiene “buscando” en la columna de H_i el primer valor de la variable con frecuencia relativa acumulada mayor que 0'75. Si algún valor de x_i tiene $H_i = 0'75$, Q_3 será la media de dicho valor y el siguiente.